Lecture 7: Linear transformations

Tuesday, September 15, 2015 9:30 AM

Admin: Reading: 4.3–4.4
Midterm

Today: Linear transformations $\mathbb{R}^n \to \mathbb{R}^m$
Goal: Understand the geometry of linear maps.

Recall:

Definition: For an $m \times n$ matrix A,

- The nullspace of A is
$$N(A) = \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = 0\}.$$
- The range (or column space) of A is
$$R(A) = \{A\vec{x} \mid \vec{x} \in \mathbb{R}^n\}.$$
$$= \text{Span(columns of } A)$$
- The rowspace of A is $R(A^T)$

Example:

Definition: The matrix A is diagonally dominant if for all rows j,
$$|a_{jj}| \geq \sum_{i \neq j} |a_{ij}| .$$

It is strictly diagonally dominant if the inequality is strict ($>$) for all rows.

Theorem: If A is strictly diagonally dominant, then
$$N(A) = \{\vec{0}\}.$$

Proof:
Let $\vec{x} \in N(A)$, so $A\vec{x} = \vec{0}$.
Let j be the coordinate so $|x_j|$ is largest.
$$0 = (A\vec{x})_j = a_{jj}x_j + \sum_{i \neq j} a_{ji}x_i$$
$$\Rightarrow |a_{jj}x_j| = \Big|\sum_{i \neq j} a_{ji}x_i\Big|$$
$$\leq \sum_{i \neq j} |a_{ji}| \cdot |x_i|$$
$$\leq \Big(\sum_{i \neq j} |a_{ji}|\Big) \cdot |x_j|$$
$$< |a_{jj}| \cdot |x_j| , \text{ unless } x_j = 0.$$

$\le \left(\sum_{i \neq j} |a_{ji}|\right) |x_j|$

$< |a_{jj}| \cdot |x_j|$, unless $x_j = 0$.

This is a contradiction unless $x_j = 0$.

$\Rightarrow x_j = 0 \Rightarrow \vec{x} = \vec{0}$

$\Rightarrow N(A) = \{\vec{0}\}$. ✓ $\square$

Question: Can you characterize the nullspaces of diagonally dominant matrices?

Example:

$$A = \begin{pmatrix} -2 & 1 & & & & 1 \\ 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & & \ddots & & \\ & & & 1 & -2 & 1 \\ 1 & & & & 1 & -2 \end{pmatrix}$$

← diagonally dominant, but not strictly so

has nullspace

$$N(A) = \mathrm{Span}(\{(1,1,1,\ldots,1)\})$$

= set of all constant vectors.

LINEAR TRANSFORMATIONS

Why matrices???

Why matrix multiplication?
Why matrix inversion?

(Why do only square matrices have inverses?)

Definition: A function $f: \mathbb{R}^n \to \mathbb{R}^m$ is linear if

$$f(\alpha \vec{u}) = \alpha \cdot f(\vec{u})$$ for all vectors $\vec{u} \in \mathbb{R}^n$ and scalars $\alpha \in \mathbb{R}$

$$f(\vec{u} + \vec{v}) = f(\vec{u}) + f(\vec{v})$$ for all $\vec{u}, \vec{v} \in \mathbb{R}^n$

Examples: For $n = m = 2$,

- $f(x,y) = (0,0)$ ✓
- f a rotation by θ: $f(x,y) = \begin{pmatrix} (\cos\theta)x - (\sin\theta)y, \\ (\sin\theta)x + (\cos\theta)y \end{pmatrix}$ ✓
- $f(x,y) = (x^2, \sin y)$ ✗ not linear
- $f(x,y) = (1+x, y)$ ✗ not linear

One more example: For a polynomial p,

e.g., $5x^2 + 3x + 2$,

let

$$f(p) = (2 + 3x) \cdot p.$$

f maps polynomials in x to polynomials in x

← a vector space

and it is a linear transformation!

f maps polynomials in x to polynomials in x
$\leftarrow$ a vector space
and it is a linear transformation!

We'll see lots more examples later (e.g., differentiation, ...)

LINEAR TRANSFORMATIONS
$\updownarrow$
MATRICES

We'll see the correspondence today for maps $\mathbb{R}^n \to \mathbb{R}^m$, and generalize it to arbitrary vector spaces next week.

Theorem 1: Let A be an $m \times n$ matrix.
Define $f: \mathbb{R}^n \to \mathbb{R}^m$ by
$$f(\vec{v}) = A\vec{v}.$$
Then f is a linear transformation.

Proof: $f(\alpha u) = A \cdot (\alpha u) = \alpha \cdot (Au) = \alpha f(u)$ ✓
$f(u+v) = A \cdot (u+v) = Au + Av = f(u) + f(v)$. ✓ □

Theorem 2: Let $f: \mathbb{R}^n \to \mathbb{R}^m$ be any linear transformation. Then there exists an $m \times n$ matrix A such that
$$f(\vec{v}) = A\vec{v}.$$

Proof: Let $e_j = (0, 0, \dots, \underset{\uparrow \text{ jth coord}}{1}, 0, \dots, 0) \in \mathbb{R}^n$.

Let
$$A = \begin{pmatrix} | & | & & | \\ f(e_1) & f(e_2) & \cdots & f(e_n) \\ | & | & & | \end{pmatrix}$$
; its jth column is $f(e_j)$.

We claim that $f(\vec{u}) = A\vec{u}$ for any $\vec{u} \in \mathbb{R}^n$. Indeed,
$$\vec{u} = (u_1, u_2, u_3, \dots, u_n) = u_1\vec{e}_1 + u_2\vec{e}_2 + \cdots + u_n\vec{e}_n$$
$$\Rightarrow f(\vec{u}) = u_1 f(\vec{e}_1) + u_2 f(\vec{e}_2) + \cdots + u_n f(\vec{e}_n)$$
by applying the linearity property repeatedly.
The right-hand side is
$$A\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = A\vec{u}. \quad ✓$$
□

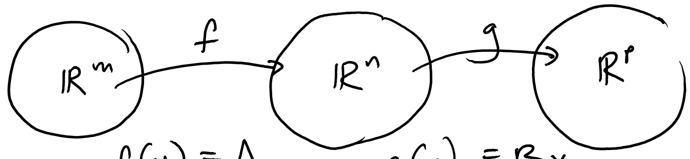
These theorems are why matrix-vector multiplication is defined the way it is.

What about matrix-matrix multiplication?

MATRIX $\longleftrightarrow$ LINEAR FUNCTION

What about matrix-matrix multiplication?

$$\boxed{\text{MATRIX MULTIPLICATION} \longleftrightarrow \text{LINEAR FUNCTION COMPOSITION}}$$



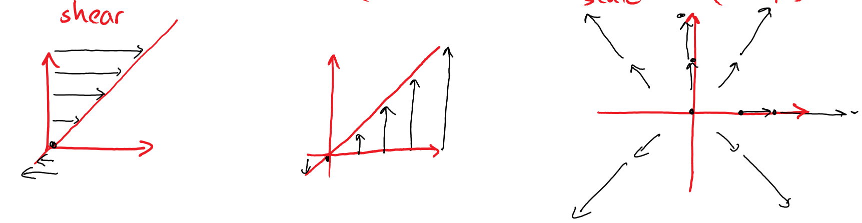
$f(u) = Au \qquad g(v) = Bv$

$$g(f(u)) = B \cdot A \cdot u$$

How to visualize matrices

I. 2×2 matrices as linear transformations

$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ shear $\qquad A^T = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \qquad S(\alpha,\beta) = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}$ scale



Claim: Any 2×2 matrix can be expressed as a product of shearing and scaling matrices, eg.,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & c \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} bc & 0 \\ 0 & ad-bc \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \frac{1}{bc} & 0 \\ 0 & \frac{1}{ac} \end{pmatrix}$$

(if $abc \neq 0$)

Proof: Starting with any 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, G.E. row operations change it to a matrix of the form either

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & x \\ 0 & 0 \end{pmatrix} \text{ or } \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

column operations

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

These final matrices are $S(0,0), S(1,1), S(1,0)$.

Row operations (scaling a row, adding one row to another) can be implemented by multiplying on the left by A, A^T or $S(\alpha,\beta)$. Column operations by right-multiplying. Therefore we can get from the end $S(0,0), S(1,1)$ or $S(1,0)$ to the beginning $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ by left- or right-multiplying by the generator matrices. ✓ □

Exercise: Generalize this claim to $n \times n$ matrices. Using row and column operations, acting on two

Exercise: Generalize this claim to $n \times n$ matrices.

Using row and column operations, acting on two coordinates at a time, any $n \times n$ matrix can be reduced to the form

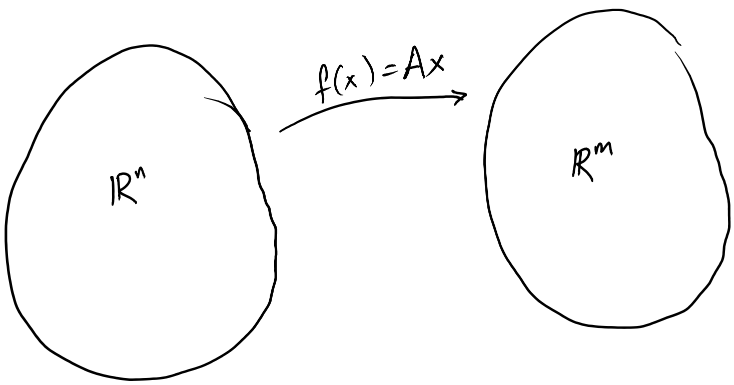
$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & 0 & \\ & & 1 & & & \\ & & & 0 & & \\ & 0 & & & \ddots & \\ & & & & & 0 \end{pmatrix}.$$

Singular-Value Decomposition (SVD)

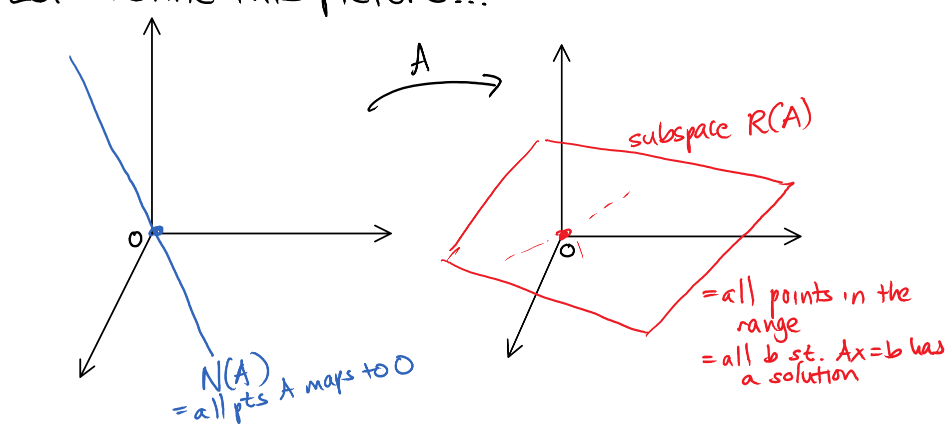
Informally: Any linear transformation can be split into:
- a rotation, followed by
- scaling vectors in or out

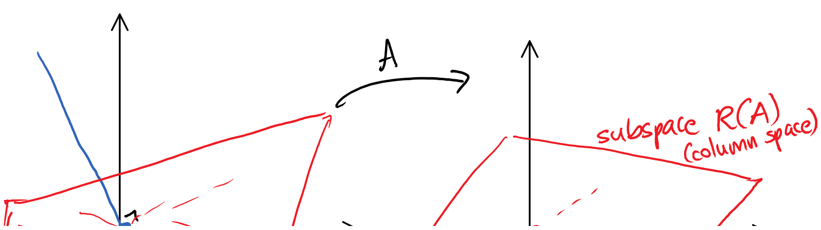
II. The geometry of linear transformations

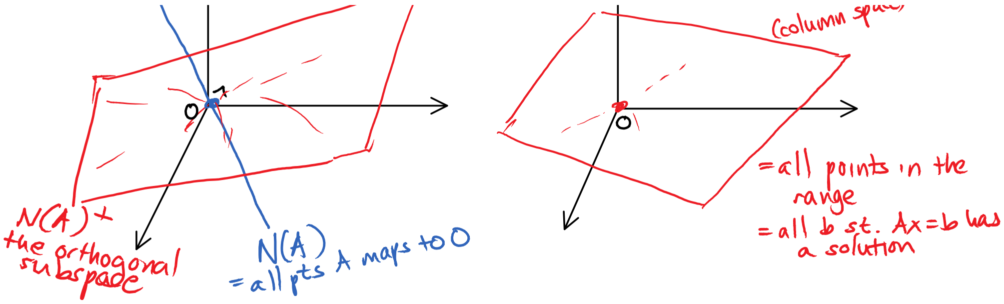
Let A be an $m \times n$ real matrix.



Let's refine this picture...

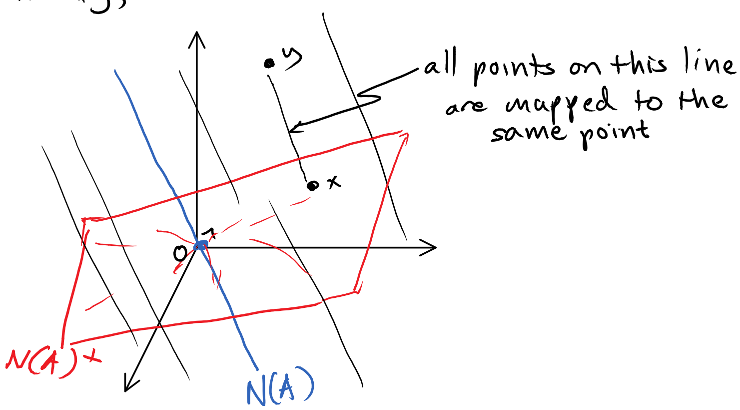






Observations

① If $\vec{x}-\vec{y} \in N(A)$, then $A\vec{x} = A\vec{y}$.
Graphically,

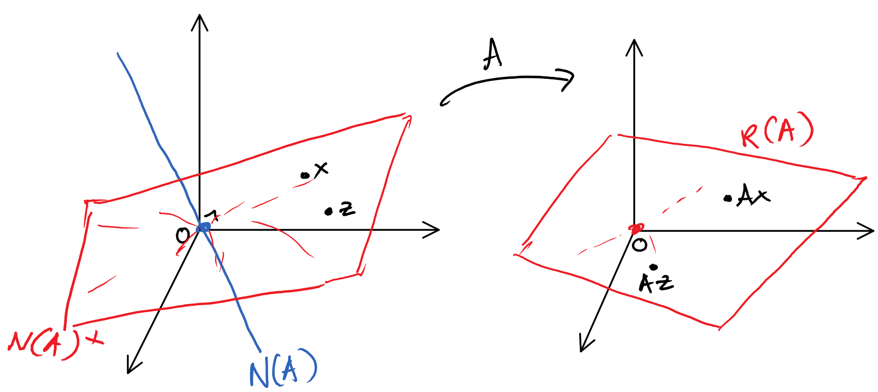


In our heads, we can thus break A into two steps:

1) First map y to x,
   i.e., take a point and move it parallel to $N(A)$ to get to $N(A)^\perp$.
   This is a **projection**; it flattens the space to $N(A)^\perp$.

2) Map $\vec{x}$ to $A\vec{x}$.

② If $\vec{x}, \vec{z}$ are distinct points in $N(A)^\perp$, then $A\vec{x} \neq A\vec{z}$.
(because if $Ax = Az$, then $x - z \in N(A)$.)



Thus A is a 1-to-1 map $N(A)^\perp \rightarrow R(A)$. It is invertible between these spaces.

This picture is not yet complete. We will prove later that between these spaces.

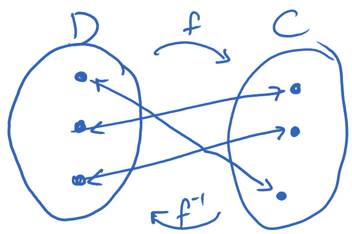
This picture is not yet complete. We will prove later that

$$N(A)^\perp = R(A^T) \leftarrow \text{rowspace of } A$$

and $\dim R(A^T) = \dim R(A)$ ← called the **rank** of A.

Matrix and function inverses

Definition: A function $f: \underset{\text{domain}}{D} \to \underset{\text{codomain}}{C}$ is **invertible** if every point in C is the image of exactly one point in D.

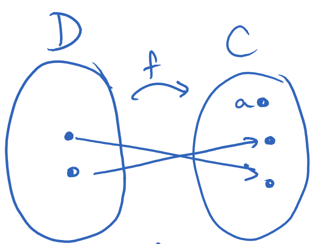


The inverse function $f^{-1}: C \to D$ takes each point in C to its unique preimage.

Equivalently, f^{-1} satisfies

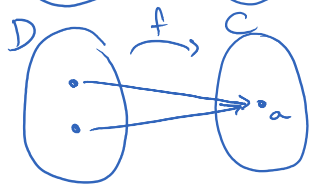
$f^{-1} \circ f$ = identity on D

$f \circ f^{-1}$ = identity on C



a isn't the image of anything

$\Rightarrow f$ not invertible



a has two preimages

$\Rightarrow f$ not invertible

Exercise: Prove that the inverse of a linear function, if it exists, is also linear.

Definition: The inverse of a matrix A is a matrix B that satisfies

BA = the identity matrix and AB = the identity matrix.

Observe:
- Not every matrix is invertible, e.g., $A = (\,0\,)$ is not invertible.

  A matrix that is not invertible is called **singular**.
- If an inverse exists, then it is unique.

Proof: Assume B and C are both inverses of A.

Consider BAC.

BAC =

$C = (BA)C \qquad B(AC) = B \quad \Rightarrow B = C$ ✓ □

Example: The 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is invertible/nonsingular if and only if $ad - bc \neq 0$. The inverse is given by

$$A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

(The proof is an exercise.)

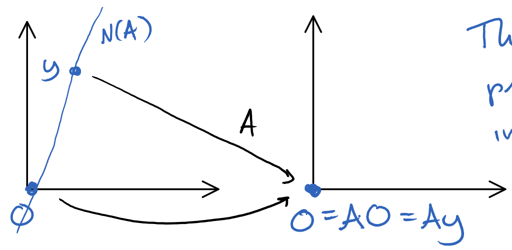
How to compute the inverse of a matrix?

Gaussian elimination, of course ...

When is a matrix invertible?

Lemma 1: If $N(A) \neq \{\vec{0}\}$, then A is not invertible.

Proof: Take $y \in N(A)$, $y \neq 0$.



Then the point 0 has two preimages, so A is not invertible. □

Lemma 2: If A is an $m \times n$ matrix with $m < n$, then $N(A) \neq \{\vec{0}\}$.

Corollary: Only square matrices can be invertible.

Proof: Let A be an $m \times n$ matrix.

$m < n$: $N(A) \neq \{0\}$ (Lemma 2)

$\Rightarrow$ A not invertible (Lemma 1) ✓

$m > n$: Assume A^{-1} exists, an $n \times m$ matrix.

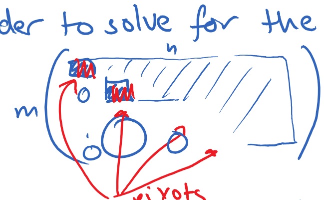
$\Rightarrow N(A^{-1}) \neq \{0\}$ (Lemma 2)

$\Rightarrow A^{-1}$ not invertible (Lemma 1)

$\Rightarrow$ contradiction, since $(A^{-1})^{-1} = A$. ✓ □

Proof of Lemma 2 ($m < n \Rightarrow N(A) \neq \{0\}$).

Applying Gaussian elimination to A in order to solve for the nullspace results in a matrix like



The point is that there are at most m pivots. Since $m < n$, there are necessarily at least $n - m \geq 1$ free variables, so the nullspace is infinite. ✓ □

Theorem: An $m \times n$ matrix A is invertible if and only if $m = n$ and $N(A) = \{\vec{0}\}$.

Proof:

($\Rightarrow$:) If A is invertible, then we have just shown that $m = n$ and $N(A) = \{0\}$. ✓

($\Leftarrow$:) Follows from:

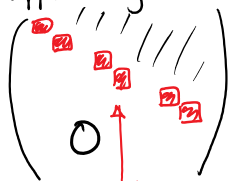
Lemma 3: For an $n \times n$ square matrix A,

$$N(A) = \{\vec{0}\} \iff R(A) = \mathbb{R}^n.$$

Proof: As usual, use Gaussian elimination.

$$A = G \; \mathcal{U} \leftarrow \text{upper triangular}$$

G: the row operations; they can be undone so G is invertible



If $\mathcal{U}$ has n pivots:

$\mathcal{U} = \text{Identity} \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$, so $A = G$ is invertible with $N(A) = \{0\}$, $R(A) = \mathbb{R}^n$.

If $\mathcal{U}$ has $r < n$ pivots:

$N(\mathcal{U}) \neq \{0\} \Rightarrow N(A) = N(\mathcal{U}) \neq \{0\}$

If $R(A) = \mathbb{R}^n$, then

$R(\mathcal{U}) = G^{-1}(R(A)) = \mathbb{R}^n$. But that's impossible; $e_n \notin R(\mathcal{U})$.

(Lemma 3) □

(Theorem) □

Note:
- Lemma 3 is false for non-square matrices.
- Lemma 3 is false in infinitely many dimensions.

Example:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ & & \vdots & & \ddots \end{pmatrix}$$

$N(A) = \text{Span}(\vec{e}_1)$, but $R(A) = $ everything.

Trivia:

HOMOGENEOUS COORDINATES & AFFINE TRANSFORMATIONS

Q: Is this a linear transformation?

$$(x, y, z) \longmapsto (x+1, y, z)$$

No. A linear transformation has to take $\vec{0}$ to $\vec{0}$, in order to satisfy linearity under multiplication:

$$f(\alpha \vec{v}) = \alpha \cdot f(\vec{v})$$
$$\Rightarrow f(\vec{0}) = \vec{0}\text{, setting } \alpha = 0.$$

Thus translations are not linear.

- "Homogeneous coordinates"

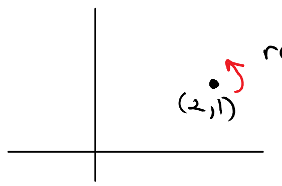
Add a new coordinate, $w = 1$. Then,

$$T: (x,y,z) \longmapsto (x+1, y, z)$$

corresponds to

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} : (x,y,z,w) \mapsto (x+w, y, z, w)$$

Example: Using homogeneous coordinates, give a 3×3 matrix for the 2D affine transformation



rotate 30 degrees about the point $(2,1)$.

Answer: Shift to the origin $(0,0)$, rotate there, and shift back.

$$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{array}{c} \begin{array}{ccc} x & y & w \end{array} \\ \begin{pmatrix} c & -s & 0 \\ s & c & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{array} \begin{array}{c} \begin{array}{ccc} x & y & w \end{array} \\ \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \end{array}$$

$c = \cos 30° = \frac{\sqrt{3}}{2}$

$s = \sin 30° = \frac{1}{2}$

$$= \begin{array}{c} \begin{array}{ccc} x & y & w \end{array} \\ \begin{pmatrix} c & -s & 2-2c+s \\ s & c & 1-c-2s \\ 0 & 0 & 1 \end{pmatrix} \end{array} = M \quad ✓$$

Sanity check: $M\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$ ✓ $\quad M\begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}$

$$M\begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} = M\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} + M\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2+c \\ 1+s \\ 1 \end{pmatrix} ✓$$

TODO

Explain how to compute A^{-1}: GE on $(A \mid I)$

b/c to solve 2 equations $Ax = b$, $Ax' = b'$, can do GE on $(A \mid b \; b')$

OUTLINE

Linear transformations $\mathbb{R}^n \to \mathbb{R}^m$ $\longleftrightarrow$ $m \times n$ matrices

matrix multiplication $\longleftrightarrow$ composition of functions

examples

only square matrices are invertible

preview of the SVD. how to compute A^{-1} by hand

COMMENTS:

- this mostly works out, except at the end showing $N(A)=\{0\} \Rightarrow A$ is invertible is kind of awkward without the notions of dimension/rank or even linear independence.
  - we haven't ~~proved~~ that subspaces in $\mathbb{R}^n$ are lines/planes/hyperplanes, i.e., that they are spans of finite sets of vectors
  - we use a basis for $\mathbb{R}^n$ and a basis expansion to prove linear transformation $\to$ matrix

SCRATCH SPACE

Theorem: An $m \times n$ matrix A is invertible if and only if $m=n$ and $N(A)=\{0\}$.

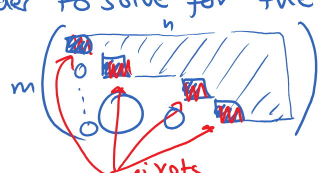
Proof:

($\Rightarrow$:) Assume A is invertible.
If $N(A) \neq \{0\}$, then there are multiple solutions to $Ax=0$, so the inverse is not defined at 0, a contradiction. $\Rightarrow N(A)=\{0\}$.

Lemma: If A is an $m \times n$ matrix with $m<n$, then $N(A) \neq \{0\}$.

The lemma implies that $m \geq n$. Since A^{-1} is an $n \times m$ matrix that is also invertible $\left((A^{-1})^{-1}=A\right)$, we have $N(A^{-1})=\{0\}$, so applying the lemma to A^{-1} gives $n \geq m$. Hence $m=n$.

Proof of lemma: Applying Gaussian elimination to A in order to solve for the nullspace results in a matrix like



The point is that there are at most m pivots. Since $m<n$, there are necessarily at least $n-m \geq 1$ free variables, so the nullspace is infinite. ✓ $\square$

($\Leftarrow$:) Assume $m=n$ and $N(A)=\{0\}$.

$$\mathrm{Ker}\begin{pmatrix} -2 & 1 & 0 & 1 \\ 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 1 & 0 & 1 & -2 \end{pmatrix} \;\; \tfrac{1}{2} \;\; \tfrac{1}{2} \longrightarrow \begin{pmatrix} -2 & 1 & 0 & 1 \\ 0 & -\frac{3}{2} & 1 & \frac{1}{2} \\ 0 & 1 & -2 & 1 \\ 0 & \frac{1}{2} & 1 & -\frac{3}{2} \end{pmatrix} \;\; \tfrac{2}{3} \;\; \tfrac{1}{3}$$

$$\begin{aligned} -2x_1 + x_2 \qquad\quad + x_4 &= 0 \\ x_1 - 2x_2 + x_3 \qquad &= 0 \\ x_2 - 2x_3 + x_4 &= 0 \\ x_1 \qquad + x_3 - 2x_4 &= 0 \end{aligned}$$

$$\to \begin{pmatrix} -2 & 1 & 0 & 1 \\ 0 & -\frac{3}{2} & 1 & \frac{1}{2} \\ 0 & 0 & -\frac{4}{3} & \frac{4}{3} \\ \cancel{0} & \cancel{0} & \cancel{\frac{4}{3}} & \cancel{-\frac{4}{3}} \end{pmatrix} \;\; -\tfrac{3}{8} \;\; -\tfrac{3}{4}$$

$$\longrightarrow \begin{pmatrix} -2 & 1 & 1 & 0 \\ 0 & \cancel{-\frac{3}{2}} & \cancel{\frac{3}{2}} & 0 \\ & 1 & -1 & \\ 0 & 0 & 1 & -1 \end{pmatrix} \xrightarrow{1} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

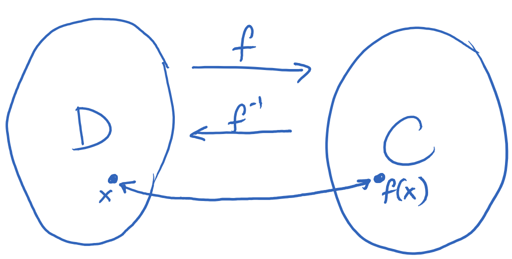
$$x_1 = x_2$$
$$x_2 = x_3$$
$$x_3 = x_4$$

$$\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \longrightarrow \begin{pmatrix} 2 & -1 & -1 \\ 0 & 3/2 & -3/2 \\ \cancel{0} & \cancel{-3/2} & \cancel{3/2} \end{pmatrix} \longrightarrow \begin{pmatrix} 2 & -1 & -1 \\ 0 & 1 & -1 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 2 & -2 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

<u>Claim</u>: $\mathrm{Range}(A) = \{(x_1, \dots, x_n) \mid \sum_{j=1}^{4} x_j = 0\}$

<u>Definition</u>: A function $f: \underset{\text{domain}}{D} \to \underset{\text{codomain}}{C}$ is **invertible** if there is a function $g: C \to D$ with $g \circ f$ the identity on D and $f \circ g$ the identity on C.



Exercise: • Prove that the inverse of a function, if it exists, is unique.

• Prove that the inverse of a linear function, if it exists, is also linear.

NOTE Linear Transformations:

p. 89: linear functions are defined (& affine fns)
examples include differentiation, trace, matrix multiplication

p. 93: the product of two matrices represents the composition of the two associated linear functions

p. 169: Range(linear function) is a subspace,
Every subspace is the range of some linear function
(namely the matrix whose columns form a finite spanning set — does it prove this exists??)

p. 238 §4.7 Linear transformations
matrix representations (with bases)
space of linear transformations is a vector space (just matrices)